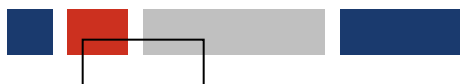




INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL



METODOLOGIA DE ESTIMAÇÃO DO IE para o 1º Trimestre de 2011



Lucília Carvalho

30 de Maio de 2011





Sumário



- Objectivo e metodologia
- Estimação das diferenças
- Estimador baseado na amostra de teste
- Estimador baseado na série temporal
- Estimador composto





Objectivo e metodologia



Objectivo

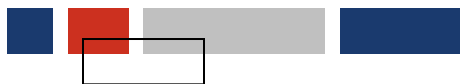
No 1º trimestre de 2011, estimar os apuramentos do IE anterior a partir dos apuramentos do IE novo para um conjunto de variáveis pré-definido

Metodologia

Envolve 4 partes:

- Estimação das diferenças entre os dois apuramentos com base na amostra de teste.
- Estimação dos apuramentos do IE antigo com base na amostra de teste
- Estimação dos apuramentos do IE com base na série temporal
- Estimador composto dos apuramentos IE





1º parte: Estimação da diferença entre inquéritos



Para cada variável **Y** estima-se a diferença entre o IE novo e o IE anterior a partir de subamostras da amostra de teste apresentadas seguidamente

Esta diferença é estimada através de um modelo linear

$$D_Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \varepsilon$$

Variável dependente

D_Y Diferença entre os dois tipos de apuramento

Variáveis explicativas

X_1 Proporção de elementos pertences à primeira rotação

X_2 Proporção de elementos entrevistados em CATI



1º parte: Estimação das diferenças entre inquéritos



Estimados os parâmetros do modelo a estimativa da diferença baseada na amostra de teste é

$$\hat{D}_Y^{AT} = \hat{\alpha}_0 + \hat{\alpha}_1 X_1(1T2011) + \hat{\alpha}_2 X_2(1T2011)$$

Vantagens do estimador

Ser baseado em amostras de teste com as seguintes propriedades:

- Têm dimensão análoga e idêntica calibração
- Têm uma parte comum e diferenças provenientes de todos os tipos de alterações que vão aumentando controladamente



2ª parte: Estimação dos apuramentos do IE anterior baseados na amostra de teste

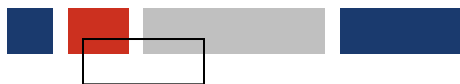


A estimativa do valor do IE anterior baseado nas amostras de teste é:

$$\hat{Y}_{IE anterior}^{AT} = \hat{Y}_{IE novo} + \hat{D}_Y^{AT}$$

Limitações do estimador

- Embora centrado tem variância grande e impossível de calcular
- Fornece estimativas pouco fiáveis para pequenos agregados cujas diferenças são calculadas com amostras de pequena dimensão
- Diferenças usadas no modelo incluem um efeito amostral que pode ser substancial e que não está presente na diferença do 1º trimestre de 2011
- Os entrevistadores do inquérito novo são em grande parte diferentes dos entrevistadores da amostra de teste



3ª parte: Estimação dos apuramentos do IE anterior baseados na série cronológica



Para combater estes efeitos foi usada outra estimativa proveniente de um estimador que, embora enviesado, tem um erro quadrático médio conhecido e mais pequeno que a variância do estimador

$\hat{Y}_{IE anterior}^{SC}$

Previsão a um passo de um modelo ARIMA ajustado à série temporal dos apuramentos trimestrais do IE para a variável **Y** com início no 1º trimestre de 2001 e final no 4º trimestre de 2010



4ª parte: Estimador combinado



As estimativas finais para o apuramento de cada variável **Y** obtêm-se através de uma combinação linear convexa dos dois estimadores apresentados

$$\hat{Y}_{IE\text{ anterior}} = \beta \hat{Y}_{IE\text{ anterior}}^{AT} + (1 - \beta) \hat{Y}_{IE\text{ anterior}}^{SC}$$

O valor de **β** varia com a distância entre os dois estimadores privilegiando-se a estimativa baseada em teste todas as vezes que os dois estimadores estão pouco afastados

Finalmente as estimativas assim calculadas são calibradas de modo a garantir que os totais coincidem com os das estimativas independentes da população e com os totais obtidos a níveis de agregação menos detalhados

4ª parte: Estimador combinado



Os valores de β foram escolhidos de acordo com a seguinte regra:

Se $Sinal(\hat{Y}^{AT} - \hat{Y}_{IE}^{4t2010}) = Sinal(\hat{Y}^{SC} - \hat{Y}_{IE}^{4t2010})$

então

$$\hat{Y}^{SC} - \sigma(\hat{Y}^{SC}) \leq \hat{Y}^{AT} \leq \hat{Y}^{SC} + \sigma(\hat{Y}^{SC}) \rightarrow \beta = 0.7$$

$$\hat{Y}^{SC} - 2\sigma(\hat{Y}^{SC}) \leq \hat{Y}^{AT} \leq \hat{Y}^{SC} + 2\sigma(\hat{Y}^{SC}) \rightarrow \beta = 0.4$$

$$\hat{Y}^{SC} - 3\sigma(\hat{Y}^{SC}) \leq \hat{Y}^{AT} \leq \hat{Y}^{SC} + 3\sigma(\hat{Y}^{SC}) \rightarrow \beta = 0.2$$

$$\hat{Y}^{AT} \leq \hat{Y}^{SC} - 3\sigma(\hat{Y}^{SC}) \text{ ou } \hat{Y}^{AT} \geq \hat{Y}^{SC} + 3\sigma(\hat{Y}^{SC}) \rightarrow \beta = 0.05$$

Se $Sinal(\hat{Y}^{AT} - \hat{Y}_{IE}^{4t2010}) \neq Sinal(\hat{Y}^{SC} - \hat{Y}_{IE}^{4t2010})$

A escolha dos valores de β ligeiramente mais baixa.